

# Blind estimation of reverberation time in classrooms and hospital wards

Paul Kendrick<sup>a</sup>, Nicola Shiers<sup>b</sup>, Rob Conetta<sup>b</sup>, Trevor J Cox<sup>a</sup>, Bridget M Shield<sup>b</sup> and Charlie Mydlarz<sup>a</sup>

<sup>a</sup>Acoustics Research Centre, University of Salford, Salford, Greater Manchester, M54WT

<sup>b</sup>Faculty of Engineering, Science and Built Environment, London South Bank University, London SE1

OAA

Corresponding author: p.kendrick@salford.ac.uk, Phone: (+44)161 295 4618

## Abstract

This paper investigates blind Reverberation Time (RT) estimation in occupied classrooms and hospital wards. Measurements are usually made while these spaces are unoccupied for logistical reasons. However, occupancy can have a significant impact on the rate of reverberant decay. Recent work has developed a Maximum Likelihood Estimation (MLE) method which utilises only passively recorded speech and music signals, this enables measurements to be made while the room is in use. In this paper the MLE method is applied to recordings made in classrooms during lessons. Classroom occupancy levels differ for each lesson, therefore a model is developed using blind estimates to predict the RT for any occupancy level to within  $\pm 0.07$ s for the mid-frequency octave bands. The model is also able to predict the effective room and per person absorption area. Ambient sound recordings were also carried out in a number of rooms in two hospitals for a week. Hospital measurements are more challenging as the occurrence of free reverberant decay is rarer than in schools and the acoustic conditions may be non-stationary. However, by gaining recordings over a period of a week, estimates can be gained within  $\pm 0.07$  s. These estimates are representative of the times when the room contains the highest acoustic absorption. In other words when curtains are drawn, there are many visitors or perhaps a window may be open.

# 1. Introduction

Measurements of room parameters such as reverberation time (RT) are usually carried out when a room is unoccupied. However it is useful to measure occupied parameters directly rather than estimate them from unoccupied measurements or perform time consuming simulations, as these can be inaccurate. For this reason a number of authors have developed blind estimation methods using reverberated speech and music as signals and facilitating in-situ, in-use measurements. For blind Reverberation Time estimation, each technique has associated advantages and disadvantages. Artificial neural networks (ANN) have been applied to the envelope spectra of speech signals [1, 2] and yield good accuracy (to within  $\pm 0.1$  s) for controlled speech signals; although the method is only semi-blind as a period of training is required. The ANN method also works for some music signals [3] but is highly sensitive to the type of music, performing more effectively with pieces containing many staccato (short) notes. Another approach is to automatically locate regions of reverberant decay and then fit a model of sound decay to those parts of the signal and from there the blind Reverberation Time is computed. Ratnam *et al.* [4, 5] utilised a stochastic model of reverberant sound decay within a maximum likelihood framework. A further enhancement to the maximum likelihood estimation method was presented by Kendrick *et al.* [3, 6-8] where a model of non-diffuse sound decay was applied, allowing the decoupling of the early and late decay regions and the automatic pruning of low dynamic range decay phases. This method has been shown in simulations to produce mid-frequency blind RT estimates to within  $\pm 0.1$  s of the RT measured using traditional methods. Only a very limited number of measurements in real environments have been carried out using the MLE method (two concert hall measurements [3]), therefore a more detailed investigation was required to validate the procedure in other real environments.

In this paper the MLE method [3, 6-8] is evaluated for use in the measurement of the occupied RT in school classrooms and hospital wards. For the classroom data, a predictive model is defined based on the Sabine equation that uses a number of occupied RT estimates combined with occupancy

level, to predict the RT over a broad range of occupancy levels. This model also provides predictions for the effective absorption in the room and the effective absorption per person. For the hospital data, a methodology is defined which allows blind RT estimates to be computed over a period of a week.

The paper is structured as follows: Section 2 introduces the blind estimation technique: Section 3 applies the algorithm to school classrooms, first performing a validation study on the effects of relative level and number of sound sources on accuracy, and then presents some results from real classroom recordings: Section 4 investigates the applicability of the method to hospital wards, with a validation study carried out first in simulated wards, then a series of estimates made using data captured over one week intervals in a real hospital wards.

## 2. Blind estimation of RT - overview

An overview of the MLE methodology is presented in Figure 1, with full details provided in reference [3]. First a length of reverberant signal is recorded, e.g. thirty minutes of speech. A linear fitting technique is then applied to short (0.01 s) sections of the signal's log-Hilbert envelope. From this, the gradient of the log-energy level is tracked and used to identify portions of the signal where the level of uninterrupted sound is continuously decaying. A stochastic model of sound decay within a room is fitted to each of these selected decay phases. The model,  $e[n]$ , is a convex sum of exponentials able to represent non-uniform decay curves:

$$e[n] = \alpha a^n + (1 - \alpha)b^n \quad (1)$$

where  $n$  is sample number,  $a$  and  $b$  represent the rate of decay of each part of the decay curve and  $\alpha$  is the mixing variable which controls the knee point between the two decay regions. The model assumes that  $e[n]$  modulates a random Gaussian variable, representing acoustic reflections. This enables a maximum likelihood approach to the estimation of the model parameters  $a$ ,  $b$  and  $\alpha$  for each decay phase which yields a large number of decay curve estimates. A data pruning algorithm

removes all decay curves with less than 25 dB of dynamic range. The remaining data is then split into a number of time windows. The MLE method makes the assumption that the decay phases whose level decreases quickest, contain the least amount of speech or background noise and are therefore most representative of free reverberant decay in the room. The fastest decreasing decay phases within a time window are used to compute a provisional blind RT estimate [3]. Due to the stochastic nature of the sound source, provisional blind RT estimates are calculated for a number of time windows and the mean of these is the blind RT estimate.

The length of the time windows used influences the accuracy. With non-stationary signals such as speech, the window length is related to the probability that the window will contain one or more regions of uninterrupted reverberant decay. For example ten, three minute sections of continuous running speech, read aloud from a book at a steady pace, produced accurate estimates ( $\pm 0.1$  s) for the octave bands 500 – 4000 Hz [3]. However real world signals will most certainly be more non-stationary and may require longer windows to ensure that free reverberant decay is present. To produce an accurate measure where there are few regions of free reverberant decay, for example during silent periods, the overall length of the recording must also be increased.

Noise in the decay curves can bias the RT estimate when Schroeder backwards integration is performed [9]. To prevent this, the later part of every decay curve estimate, from -25 dB onwards, is replaced by a decaying exponential with a decay rate estimated from the initial region of the decay curve (from -5 dB to -25 dB). The Reverberation Time ( $RT_{20}$ ) is calculated from the gradient of a least squares best fit line to the log energy of the backwards integrated decay [9] from -5 to -25 dB [10].

### **3. In-situ measurement of RT in classrooms**

#### **3.1 Introduction**

Classrooms offer an interesting case study for occupied measurement and estimation of room acoustic parameters. The acoustic conditions of classrooms can change substantially between

unoccupied and occupied states, and the presence of pupil noise provides a challenge when carrying out measurements in occupied conditions. First a series of validation experiments (section 3.2) in controlled conditions were carried out to gauge the accuracy of the blind RT estimates using the MLE method. Then, using insights from the validation study the algorithm was evaluated using a series of recordings in real classrooms (section 3.3).

## 3.2 Validation Study

Five lecture rooms within the University of Salford were selected to represent a range of classroom types with the  $RT_{20}$  ranging from around 0.4 s to 1.7 s. In each of these rooms, six acoustic impulse responses were measured using three source locations and two receiver locations. Measurements were carried out using a Genelec 7050A speaker, a B&K reference microphone and a Norsonic pre-amplifier. The rooms were excited by a swept sine wave and the impulse response extracted using WinMLS software [11].

The octave band  $RT_{20}$  for all six impulse responses was averaged and is presented in Figure 2, with the 95% confidence limits plotted as error bars for each measurement. On average for the mid-frequencies (250-4000 Hz) the confidence limits are about  $\pm 0.03$  s, the worst case 95% confidence limits are  $\pm 0.07$  s.

Source locations for a typical classroom measurement are indicated in Figure 3. One source was placed at the front of the classroom at a height of 1.6m (S1), to simulate the teacher and two at either side of the room at a height of 1.25m (S2 and S3) to simulate competing sound sources, such as noisy pupils. Receiver locations were placed either side of the room (R1 and R2). The measured impulse responses (S1R1, S1R2, S2R1, S2R2, S3R1 and S3R2) were convolved with anechoic signals before mixing to form different scenarios for the blind estimation algorithm, as described in the following section. This enabled quick and detailed investigation into the effect of parameters such as noise source type, level and location on the estimation accuracy.

### 3.2.1 Methodology

A series of thirty minute anechoic recordings of running speech were made. Four male native English speakers each read a monologue, selected arbitrarily from a number of books. One of these recordings was used to simulate the teacher speaking while up to six recordings of nine different book chapters (acoustic reference books were used) were mixed together to create competing sound sources with varying numbers of talkers. Two different single talker files and two different multi-talker (three different speakers in each) files were created where each source in each mix was normalised to have equal power. The blind estimation algorithm was applied to the three mixtures described in Table 1.

The reverberant levels of the teacher and competing noise sources were first normalised to have equal  $L_{eq}$  within the octave band being analysed. Then for each mixture, the noise signal was attenuated prior to mixing so that for each mixture there were eight different signal to noise ratios, of 5, 10, 15, 20, 25, 30, 40 and 50 dB, plus a noise free situation with only the teacher speaking.

### 3.2.2 Validation study results – noise free

The blind estimation was first applied to the noise free cases, where only the teacher sound source was present. A single measurement position was used as this is a realistic practical constraint on an occupied measurement. A blind RT estimation was calculated for each octave band separately. The found decay phases were collected into ten groups so, on average, a decay curve was produced for every three minutes of the recording. A provisional blind  $RT_{20}$  estimate was produced for each three minute section and the mean value over all sections used as the blind  $RT_{20}$  estimate. The error is quantified as the difference between the blind  $RT_{20}$  estimate and the mean  $RT_{20}$  computed from the six impulse responses. The errors, together with the 95% confidence limits, are presented in Figure 4. For the octave bands 500 – 8000 Hz the error magnitude is generally less than 0.1 s. However, the result at 250 Hz shows a significant increase in error, especially for the more reverberant room. This is because the lower frequency bands tend to be more reverberant due to less absorption at low

frequencies, and thus the dynamic range of the decay phases captured by the segmentation algorithm is reduced. The error is generally positive, which suggests a positive bias error, but this is generally less than 0.1 s for the frequencies 500-8000 Hz; the bias error will be studied further in section 3.2.4.

### 3.2.3 Validation study results – the effect of noise

Figures 5 and 6 show the effect on estimation accuracy of adding additional noise sources to the simulations and varying the signal-to-noise ratio. The largest error across all rooms is shown. Figure 5 shows the effect of varying the signal-to-noise ratio with one additional talker while Figure 6 shows the effect with six additional talkers (three at each of two speaker locations).

For a single competing talker (Figure 5) the accuracy begins to significantly deteriorate when the signal-to-noise ratio is equal to and less than 15 dB. The two talker competing noise mixture demonstrated similar accuracy (not plotted), the 2 kHz and 4 kHz octave bands being marginally worse with additional talkers and the 500 Hz and 1 kHz bands being slightly better. With the six talker babble, the performance is generally worse for all octave bands with significant deterioration for the 500 Hz, 2 kHz, 4 kHz bands (beginning at the 25 dB signal-to-noise ratio), and less significant deterioration at 1kHz and 8 kHz when compared with the one and two talker cases.

Previous studies [3] have analysed the performance of the MLE algorithm in the 1 kHz octave band, varying signal-to-noise ratios using white noise and speech as the noise and source signals. This showed increased error when signal-to-noise ratios were decreased beyond 25 dB, as in the presence of a high noise floor, the decay rate of the selected decay phases will be biased. Figures 5 and 6 show that the performance is less affected when the noise source is highly non-stationary. Non-stationary noise sources, for instance when a single extra talker is present, do not interrupt every instance of reverberant decay produced by the main sound source. In fact in some cases the noise source may increase the number of valid decay phases in the recording possibly increasing the accuracy.

### 3.2.4 Error analysis

It is important to be able to quantify the possible error in any blind estimate to infer the accuracy of the algorithm in real environments. There are a number of sources of error that may impact on the accuracy of the blind  $RT_{20}$  estimates. Random errors can be identified by multiple measurements and reduced by averaging, but bias errors are more difficult to account for or detect. To better understand the problem the random error is compared with the bias error. This is carried out for all five rooms for the frequencies from 250 Hz to 8 kHz for the one, two and six talker cases and for all signal-to-noise ratios. The blind  $RT_{20}$  estimate is computed by calculating the mean provisional  $RT_{20}$  over all ten windows. The standard error for the blind  $RT_{20}$  estimate provides an estimation of the random error and from this the 95% confidence limits are computed using the following formula;

$$CL_{95\%} = 1.96 \left( \frac{\sigma}{\sqrt{N}} \right) \quad (2)$$

where  $\sigma$  is the sample standard deviation and  $N$  is the number of samples (10). The bias error is the difference between the blind  $RT_{20}$  estimate and the mean  $RT_{20}$  calculated directly from the room impulse responses (Figure 2). Figures 7 and 8 plot the random error (95% confidence limits) verses the magnitude of the bias error for all the results. These plots show that for estimates with bias errors ( $<0.1$  s) there is no or little correlation between the bias and random parts of the error, but for higher error cases there is a stronger relationship. The reason for this is that large errors are caused by reverberant decays being interrupted by other sounds. This in turn causes the decay rate estimates to be positively biased, but as the level and decay rates of these interrupting sounds is highly variable, the variability of the provisional  $RT_{20}$  also increases. Conveniently, this means the random error can be used as an indicator of a bias error to reject data during blind estimation. For all estimates with 95% confidence limits less than 0.1 s, 88% of these estimates also have a bias error that is less than 0.1 s.



### 3.2.5 Validation study, implications towards measurements in occupied classrooms

The soundscape within a classroom varies from whole-class instruction (one talker), to group work where many voices are present simultaneously. The validation study evaluated a large number of cases (3 different numbers of talkers mixed at 7 different signal-to-noise ratios) which represent a range of cases that may occur in the classroom. In order to apply what has been learnt to real classroom recordings the density of regions of free reverberant decay and the length of the recordings must be considered.

In the validation study 30 minute recordings were made and the number of suitable decay phases per minute was found to be around twelve (for 1kHz octave band and averaged over all mixtures). These decay curves were grouped into ten groups containing around 36 decay phases in each, and the average provisional  $RT_{20}$  over the ten groups used as the blind  $RT_{20}$  estimate. In real classroom recordings, which is covered in Section 3.3, it was found that the average number of suitable decay phases per minute varied between lessons from six per minute to as low as one per minute. A lesson is usually about 50-60 minutes long so lessons producing six suitable decay phases per minute should produce blind  $RT_{20}$  estimates with similar accuracy to the validation study. After grouping the decay phases into ten groups there should be also be around 36 decay phases in each group. There will be lessons that do not produce sufficient regions of free reverberant decay to produce accurate estimates. The 95% confidence limits, which are computed using only the blind estimates, can be used to indentify cases with large bias errors. Therefore blind  $RT_{20}$  estimates with 95% confidence limits greater than 0.1 s are rejected.

The density of the suitable decay phases varies over individual lessons, but the algorithm deals with this automatically as selected decay phases are grouped into windows after they have been identified. For example, 36 decay phases may be extracted from a single ten minute section while the next 36 suitable decay phases could be collected from the next twenty minutes.

### 3.3 Results from real classroom measurements

Three classrooms were selected, from data collected by the ISESS project [12, 13] to represent a range of typical reverberant environments; rooms A, B and C, used for Maths, English and Science respectively. The unoccupied (except for the researcher) octave band  $RT_{20}$  was measured using three balloon bursts and a Norsonic 140 sound level meter (SLM) using three source positions and one receiver position. Balloon bursts were used, as a convenient measurement method was required [10]. The mean  $RT_{20}$  for three measurements and the 95% confidence limits for the three classrooms are presented in Figure 9. Note that the worst case over the mid frequency bands (500-4000 Hz) is  $\pm 0.07s$ .

For the blind reverberation time measurements, five lessons in each classroom were recorded using the same SLM with 24 bit resolution and a sampling frequency of 44.1 kHz. Each recorded lesson was about 50 minutes long. The number of persons within the room in each lesson was recorded including the teacher and researcher, see Table 2 which also shows the volume of each classroom.

The MLE algorithm was applied to the octave band filtered data. The decay phases from each lesson were segmented into eight windows and a provisional  $RT_{20}$  calculated for each window. Once again the blind  $RT_{20}$  estimate is calculated as the mean provisional  $RT_{20}$ , and as demonstrated in the validation study, data sets are where the 95% confidence limits were greater than  $\pm 0.1 s$  were rejected.

Throughout the school day there can be significant variability in occupancy which may result in a large change in the Reverberation. If the blind  $RT_{20}$  estimates were made over a range of occupancies, it is possible to predict the  $RT_{20}$  for any given occupancy. Therefore a model is defined based on Sabine's equation [14],

$$RT = \frac{0.161V}{A_R + nA_P} \quad (3)$$

Where  $A_R$  is the total absorption area of the empty room,  $A_P$  is the absorption area per person,  $n$  is the number of people present and the room volume is  $V$  (air absorption can be neglected for small rooms). By minimising the total squared error between a set of blind RT estimates for a range of occupancies and the predicted RT expressed in eq. (3), least mean square estimates of  $A_R$  and  $A_P$  can be made (assuming a known room volume). This two parameter optimisation is carried out using a combined grid search and a constrained optimisation [15] approach.

To calculate the 95% confidence limits on the predicted RT,  $A_P$  and  $A_R$ , a bootstrap method is used. The bootstrap method works by creating many different sub-groups from the blind RT estimates, each group is created by randomly selecting a subset from all blind RT estimates. Five hundred groups are created, and the optimisation is carried out for each group yielding predicted values of RT,  $A_P$  and  $A_R$  for each group. The variance of the resultant parameters is used to compute the 95% confidence limits for each.

The unoccupied measurement set only contained three samples (three balloon bursts), while the occupied measurements contained many more. This meant that the line fitting procedure gave lower importance to the unoccupied measurement, despite its lower variance. To give the unoccupied measurement equal importance to the blind estimates, a new unoccupied sample set is generated using a Gaussian random number generator where the mean and variance are defined by the original three unoccupied measurements. The size of this new data set is defined by the mean sample size for the blind RT<sub>20</sub> measurements. When only blind estimates are available this subset resizing is not required.

The predicted RT<sub>20</sub> verses occupancy level is presented in Figures 10, 11 and 12 as a solid line, dashed lines indicate the 95% confidence limits on the prediction. The individual occupied RT<sub>20</sub> measurements are also presented, estimates with confidence limits less than  $\pm 0.1$  s are presented as diamonds while data with confidence limits greater than  $\pm 0.1$  s, which are not used in the optimisation, are presented as squares.

To provide an indication of the accuracy of this method, the maximum value for the 95% confidence limits over the whole range of occupancy levels is presented in Table 3. This shows that for the octave bands 500 to 4000 Hz, the worst case accuracy of the method is  $\pm 0.07$  s. The error is increased at 250 Hz where the worst case over all the measurements is  $\pm 0.15$  s. This is a result of the lower signal-to-noise ratio in the lower octave bands.

Figures 13 and 14 present the estimates of  $A_P$  and  $A_R$  calculated using the model in Eq. (3). As discussed above, the 95% confidence limits for each parameter are calculated using a Bootstrap estimation of the parameter variance. Figure 13 shows that Room B contains the most amount of absorbing material and  $A_R$  is largest, while Room A classroom contains the least.

Figure 14 compares the blind estimated  $A_P$  parameters with a number of examples from the literature [16-19]. In rooms with high values of  $A_R$ , the error in the  $A_P$  estimate is greater because each person has less effect on the room absorption and so the problem is ill-conditioned. This can be seen in Figure 14 where in the English classroom the 95% confidence limits are significantly greater than in the Science and Maths classrooms.  $A_P$  represents the change in absorption area per person, and depends not only on the absorption introduced per person, but also on the chair and the closeness of the seating, therefore this value will be unique for each room. Figure 14 shows that the results for the Rooms A and C generally fall within the spread of parameter values seen in the literature. The results from the Room B class seems to deviate from the range of values seen in the literature, but due to the ill-conditioned nature of the problem and as shown by the error bars, the expected error for this dataset is high. For  $A_P$ , the average accuracies for rooms A, B and C are  $\pm 0.13$ ,  $0.25$  and  $0.1 \text{ m}^2$  respectively. For  $A_R$ , the average accuracies for rooms A, B and C rooms are  $\pm 0.6$ ,  $2.4$  and  $1.16 \text{ m}^2$  respectively. To increase the accuracy of  $A_P$  a greater number of measurements over the range of occupancy levels are required.

## 4. In-situ measurement of RT in hospital wards

### 4.1 Introduction

This section looks at the applicability of the MLE method to the in-use, blind estimation of RT in hospital wards where many of the sound sources are different to those tested previously. Hospital measurements are particularly challenging as the occurrence of free reverberant decay is rarer than in the schools recordings and the acoustic conditions may be non-stationary. A series of validation studies were first carried out in simulated hospital wards, used for training purposes, with known RTs; the results of which were used to inform on the correct methodology and indicate the expected accuracy when measurements are carried out in real wards.

### 4.2 Equipment and methodology

When making measurements over long periods of time, a single continuous recording would not be permitted in an occupied hospital ward for reasons of privacy and confidentiality. Therefore a data reduction process is introduced [20]. A Norsonic 140 Class 1 Sound Level Meter (SLM) was used in occupied hospital wards to record discrete sound files whenever the A-weighted noise level ( $L_{A_{fmax}}$ ) exceeded 70 dB(A). This provides the MLE method with an efficient data reduction system where a day's data would consist of 100-400 'trigger files' each between 6 s and 10 s long. These trigger files contained speech and impulsive sounds such as of rubbish bins closing, doors banging, dropped objects etc. Impulsive sounds are good for the MLE method as these signals have a very short duration and occur sporadically, so the resulting reverberant decay is less likely to be interrupted by subsequent impulsive sounds compared with running speech.

All sounds were recorded with 24 bit accuracy at a 12 kHz sampling frequency. This enabled blind  $RT_{20}$  estimations up to the 4 kHz octave band to be computed. The high bit depth was required to ensure maximal dynamic range in the recording. The gain settings on the SLM were set to ensure no clipping took place while maintaining sufficient dynamic range for the blind estimation method.

### 4.3 Validation study

The validation study was carried out in the clinical skills laboratories at London South Bank University (Figure 15). This room is used for clinical training purposes and standard hospital furniture is used, including beds with rails, bed tables, dustbins, sinks, privacy curtains, etc. Within this environment it is easy to create a soundscape similar to that of a hospital ward.

Two 60-minute recordings were made within a simulated ward. One measurement was made with all the curtains surrounding the beds drawn and another without the curtains, as shown in Figure 15. The soundscape was created by two people simulating noise that would be comparable to that found in a hospital ward, based on experience of observing ward sounds. Noise included conversation, moving of furniture and bed rails, use of rubbish bins and sinks, dropping objects and opening and closing of doors. During each measurement period approximately 150 trigger files were created, which in total were approximately 20 minutes in length. Care was taken not to locate the SLM too close to the sound sources to ensure that the reverberant energy was sufficient compared with the direct sound.

#### 4.3.1 RT Measurements

Using thick latex balloons as an impulsive noise source, a series of six impulse response measurements were carried out in each condition of the room, with three source and two receiver positions. The results are shown in Figure 16. The effect of drawing the curtains around the beds can be seen to reduce the  $RT_{20}$  by between 0.1 and 0.3 s.

The MLE method required modification to utilise the trigger files. The time stamped trigger files were split into a number of groups, these groups are equivalent to the time windows used in the schools measurements. The fastest decaying curve in each group was used to produce a single provisional  $RT_{20}$  for each group. The blind  $RT_{20}$  estimate is the mean provisional  $RT_{20}$  over a number of groups. For this validation study, the data was split into five groups and thus an estimate was produced, on average, every twelve minutes.

### 4.3.2 Blind estimation validation results

Figure 17 presents the data from the validation study, showing the estimations with and without the curtains drawn. The MLE method has been able to detect that there is less reverberation when the curtains are drawn round the beds. However, the blind  $RT_{20}$  estimated for the both open and closed curtain cases show 95% confidence limits which are generally greater than  $\pm 0.1$  s. Therefore the validation study indicates that one hour's worth of recording is unsuitable for producing accurate blind estimates of  $RT_{20}$  in a hospital ward.

The reason for this is due to the small number of decay phases with free reverberant decay collected. There are on average 1.4 suitable decay phases per minute in this validation data set, over eight times less than in the validation study for the schools. This may be partly due to the method of data collection; data is collected only when  $L_{A_{fmax}}$  exceeds 70 dB(A). Low level sounds, which may well contain regions of free reverberant decay, are lost. There are two ways to compensate for this; the recording threshold could be reduced or the recording length increased.

### 4.3.3 Validation study implications

It is clear from the validation study that one hour is insufficient time to gain accurate blind  $RT$  estimates in hospital wards using the MLE method. The hospital validation study yielded on average 1.4 examples of free reverberant decay per minute which was eight times lower than the density in the schools data. However, when real hospital recordings were examined in the proceeding section, this density was about 8 times lower still, where on average 0.18 examples of free reverberant decay per minute were present. The schools validation study suggested that to gain a good estimate, each window should contain at least 36 samples of free reverberant decay to produce one provisional  $RT_{20}$ , and at least 10 of these windows are required to gain an accurate blind  $RT_{20}$  estimate. Therefore in a hospital the window length required is at least 4 hours, and at least forty hours of data are required for an accurate blind  $RT_{20}$  estimate.

#### 4.4 Estimation of RT in real hospital wards

The SLM was located in a number of occupied wards in two major UK hospitals, for a period of seven days in most locations. The locations measured included multiple occupant bays, single rooms and nurse stations. This provided a large data set from which to estimate the reverberation time. The captured data was segmented into day (6am-6pm) and night-time (6pm-6am). This data segmentation was carried out for two reasons. Firstly, as the number of night-time trigger files was up to seven times lower than during the day the inclusion of these results could reduce the accuracy. Secondly, should there be sufficient data, it would be interesting to compare the acoustic conditions between day and night, as differing conditions such as occupancy level (no visitors during the night) and the drawing of curtains around the beds may have a significant impact on the RT.

Blind  $RT_{20}$  estimates are computed from the daytime noise data by grouping a week's worth of recordings into about twenty, four hour windows. The mean and 95% confidence limits of these blind  $RT_{20}$  estimates are then computed. Figures 18 and 19 show a selection of the blindly estimated results from nine hospital rooms in the two hospitals. The worst case 95% confidence limits for the 500 – 4000 Hz octave bands, for all daytime estimates is  $\pm 0.07$ s. The 500 – 4000 Hz octave band blindly estimated  $RT_{20}$ , for these wards, all fall between 0.25 s and 0.7 s. The 4 bed bay measurement in Figure 18 shows confidence limits which are greater than  $\pm 0.1$  s for 250 Hz and therefore are not reliable. In fact the worst case confidence limits for the 250Hz octave band for all daytime data was  $\pm 0.17$  s, which suggests the method is less useful for lower frequencies. Figure 19 shows that the rooms in Hospital 2 are generally more reverberant than hospital 1. For example, in the mid frequency bands, the 4 bed bay in hospital 1 has on average a blind  $RT_{20}$  estimate of 0.3 s while the 4 bed bay in Hospital 2 has a blind  $RT_{20}$  estimate on average of 0.5 s. The lower  $RT_{20}$  in hospital 1 for the 4 bed bays, is due to the acoustic ceiling tiles installed in hospital 1 compared with hospital 2 where all rooms utilised a ceiling heating system with perforated metal tiles covering heating pipes and a mineral wool backing.



#### 4.4.1 In-situ measurement of RT in a four bed hospital bay before and after replacement of ceiling tiles

A series of recordings were able to be carried before and after the replacement of ceiling tiles in one room in Hospital 1. The original reflective, plaster tiles were replaced by acoustic ceiling tiles. Figure 20 shows a reduction of the estimated reverberation time after the ceiling tiles were replaced by the more absorbent type. The data shows an estimated reduction in  $RT_{20}$  in the frequency bands 1000 Hz – 4000Hz by over 0.1 s. At 500 Hz the estimated  $RT_{20}$  reduction is much less and at 250Hz the 95% confidence limits are  $> 0.1s$  and therefore are not accurate. The data for the post ceiling replacement has very low error bars compared with all the other measurements, because measurements were actually carried out over two weeks.

#### 4.4.2 Comparison of day and night time measurements

The night time data, in general, had larger variability and often had 95% confidence limits  $> 0.1 s$  which precluded it from being compared with the daytime data. This meant comparison between day and night time was not possible in most cases, but for some rooms there was sufficient data. Figure 21 compares the day and night time blind  $RT_{20}$  estimates for three rooms. It can be seen that there is little difference between day and night time measurements for the 6 and 12 bed bays, but for the 4 bed bay the day time blind  $RT_{20}$  estimate is generally about 0.05 s lower and the difference has been shown to be statistically significant for the octave band 2 kHz and 4kHz (computed using a t-test  $p<0.01$ ). This difference suggests that the level of absorption present during the night is slightly lower than during the day. This is not unexpected as curtains are not drawn during the night, there are no visitors, and there is less clinical activity.

Within the hospital environment, the acoustic conditions are constantly changing due to the occupancy level, use of privacy curtains, open windows and many other aspects. Unlike the school measurement case, where for a period of a lesson the acoustic conditions can be expected to be stationary and the occupancy known, hospital measurements are more challenging. There is no easy

way to record the occupancy or other affecting conditions dynamically for each trigger file.

Therefore the blind estimates are snapshots of the acoustic conditions. If the acoustic conditions are stable then this blind  $RT_{20}$  estimate is a good representation of the reverberation in the occupied room. However if the conditions are highly variable, then the estimate can only be considered to be representative of the highest occupancy/highest absorption conditions. This is because the MLE method searches for the fastest decaying region over a period of four hours, and there may well be multiple changes in acoustic conditions over this period. Therefore, when quoting this data it should be made clear that the estimates represent periods of time when the absorption within the space is highest.

However, where night time estimates were available, these did not generally show statistically significant differences between the day time measures, and when they did, the difference was small (less than the difference limens for RT). At night, activity is minimal and the acoustic conditions will not vary significantly. The absorption is lowest during night time as the occupancy will be minimal and the curtains are not drawn around the beds. Therefore, as day time measures are snap shots of when the room contains the most absorption, comparing the day and night results gives an indication of by how much the acoustic conditions vary. As day and night time measures were similar this implies that the acoustics conditions in hospital wards are fairly stationary and the RT does not vary greatly with time.

## 5. Conclusions

This paper has investigated the applicability of blind reverberation time estimation method in school classrooms and hospital wards. For the classroom data, by recording multiple lessons, a model can be defined which predicts the  $RT_{20}$  for a range of occupancies to within  $\pm 0.07$ s (worst case) for the frequency bands 500 Hz to 4 kHz. Measurements of balloon bursts and swept sine excitation were shown to produce similar levels of accuracy, where the worst case produces an error of  $\pm 0.07$  s. The

1 predictive model is also able to estimate the effective absorption area due to absorbing surfaces in  
2 the room and the effective absorption area per person. The accuracy of these predictions varied  
3  
4 depending on the range of occupancies measured and the effect of occupancy on the rate of  
5  
6 reverberant decay. For Classrooms A and C, where the effect of additional persons on the rate of  
7  
8 reverberant decay was significant, the accuracy of the absorption area per person was  $\pm 0.13\text{m}^2$ ; the  
9  
10 accuracy decreased significantly to  $\pm 0.26\text{ m}^2$  for the classroom B where the effect of occupancy on  
11  
12 the rate of reverberant decay was low. The estimated absorption area per person was comparable  
13  
14 to that measured in similar scenarios by other authors.  
15  
16  
17  
18

19 For both scenarios a validation study was first carried out to gauge the accuracy of the method and  
20  
21 to inform on the best practice for the MLE method in that environment. These showed that when  
22  
23 many competing sounds sources are present the error is significantly increased as there are  
24  
25 insufficient instances of uninterrupted free reverberant decay. This error has a random component  
26  
27 that is easily accounted for by ensemble averaging, but there is also a systematic error, which cannot  
28  
29 be accounted for, caused by the interruption of free reverberant decay by additional sounds.  
30  
31  
32

33 However, the increase in the bias error is conveniently accompanied by a significant increase in the  
34  
35 random error. This is because the sounds that interrupt the free reverberant decay cause a large  
36  
37 increase in the random error also. It was shown that 88% of the blind  $RT_{20}$  estimates with 95%  
38  
39 confidence limits less than  $\pm 0.1\text{ s}$  were within  $0.1\text{ s}$  of the  $RT_{20}$  calculated from repeated balloon  
40  
41 bursts. This fact allows the data to be pruned blindly and estimates that are biased removed.  
42  
43  
44  
45

46 A study was also carried out to investigate the application of the MLE method to recordings of in-use  
47  
48 hospital wards. In this case a validation study was carried out in a simulated hospital ward where  
49  
50 one hour's worth of data was collected in the form of trigger files, recorded every time the  $L_{A\text{fmax}}$   
51  
52 exceeded  $70\text{ dB(A)}$ . The validation study showed that the number of periods of free, uninterrupted  
53  
54 reverberant decay per minute was around eight times lower than in the classroom validation study.  
55  
56 This meant significantly longer recordings were required. Recordings were then made at two  
57  
58  
59  
60  
61  
62  
63  
64  
65

hospitals in a number of rooms over for seven day periods. The data was segmented into day and night and a decay curve estimate was produced roughly every four hours. For daytime blind  $RT_{20}$  estimations, the worst case 95% confidence limits indicated a maximum error of  $\pm 0.07$  s in the octave bands 500 – 4000 Hz. Night time blind  $RT_{20}$  estimations showed increased variability due to the lower number of trigger sound files collected, but in a number of cases estimates were possible. Comparison of day and night blind  $RT_{20}$  estimations showed the difference was at most 0.05 s.

When interpreting these results it must be considered that the blind estimate is a snap-shot of the reverberation time in non-stationary acoustic conditions. The MLE method searches for regions of decay that decrease the quickest, and therefore the estimate that is reported will be representative of the conditions where absorption is highest (highest occupancy, curtains drawn, etc.). Night-time conditions are more stable due to low activity and therefore these estimates are representative of lowest absorption conditions. Comparing day and night blind  $RT_{20}$  estimates therefore provides an indication of by how much the acoustic conditions vary. The small difference that could be detected suggests that the  $RT_{20}$  does not vary by more than 0.05 s. However, to confirm this a longer measurement period is required, to ensure that reliable night time measurements could be made. Additionally if measurement periods could be extended significantly to a period of a month or more, it may be possible to decrease the window size and use the data to provide an indication of how the  $RT_{20}$  changes over a day.

These investigations have shown that the blind  $RT_{20}$ , MLE estimation method demonstrates similar accuracy to standard measurement methods such as balloon bursts or swept-sine measurements. The resulting occupied blind  $RT_{20}$  estimates are arguably more representative of the acoustic conditions than unoccupied  $RT_{20}$  measurements made using traditional methods. The experience of this study has highlighted that with any blind measurement method it is not possible to control the quality of the data. The strength of the blind algorithm is in the ability to blindly discard low quality data. When using the MLE method there are two main aspects to consider to achieve accurate

results. Firstly, the window length over which a single provisional  $RT_{20}$  estimate is calculated: a rule of thumb suggested by this study is that the window length should be set such that at least 36 periods of free reverberant decay are present. Secondly, at least ten of these windows are required to produce the blind  $RT_{20}$  estimate. Despite this, some of data will still be poor, but by rejecting all data where the 95% confidence limits exceeds  $\pm 0.1$  s ensures that when an estimation is available, it is accurate.

## Acknowledgements

The work on schools was partly funded by EPSRC (Engineering and Physical Sciences Research Council) (EP/G009791/1). The work on hospitals was funded by Arup Global Healthcare and EPSRC.

## References

- [1] T. J. Cox, F. Li, and P. Darlington: Speech transmission index from running speech: A neural network approach, *J Acoust Soc Am* **113**, (2003).
- [2] F. F. Li, Extracting room acoustic parameters from received speech signals using artificial neural networks, P.h.D. thesis, Acoustics Research Centre, University of Salford, 2002.
- [3] P. Kendrick, Blind estimation of room acoustic parameters from speech and music, P.h.D. thesis, Acoustic Research Centre, University of Salford, Salford, 2009.
- [4] R. Ratnam, D. L. Jones, B. C. Wheeler, W. D. O. B. Jr, C. R. Lansing, and A. S. Feng: Blind estimation of reverberation time, *J Acoust Soc Am*, **114**, (5), (2003), 2877-92.
- [5] R. Ratnam, D. L. Jones, and J. William D. O'Brien: Fast Algorithms for Blind Estimation of Reverberation Time, *IEEE Signal Processing Letters*, **11**, (6), (2004).
- [6] P. Kendrick, T. J. Cox, F. L. Francis, Y. Zhang, and J. Chambers, Blind estimation of clarity, centre time and deutlichkeit from speech and music signals, in 19th International Congress on Acoustics, Madrid, 2-7 September 2007.
- [7] P. Kendrick, T. J. Cox, F. F. Li, Y. Zhang, and J. A. Chambers: Monaural room acoustic parameters from music and speech, *J Acoust Soc Am*, **124**, (1), (2008).
- [8] P. Kendrick, F. F. Li, T. J. Cox, Y. Zhang, and J. A. Chambers: Blind Estimation of Reverberation Parameters for Non-Diffuse Rooms, *Acta Acust Acust*, **93**, (2007), 760 -770.
- [9] M. R. Schroeder: New Method of Measuring Reverberation Time, *J Acoust Soc Am* **37**, (1965), 409.
- [10] *Acoustics - Measurement of the reverberation time of rooms with reference to other acoustical parameters* ISO Standard ISO 3382:1997, 1997.
- [11] M. S. Development, "WinMLS 2004," 2004.
- [12] L. S. B. University, Project - Identifying a Sound Environment for Secondary Schools (ISESS) <http://www.lsbu.ac.uk/isess/>, last accessed 17th Aug 2011.
- [13] R. Conetta, B. M. Shield, T. J. Cox, J. E. Dockrell, and D. Connolly, A preliminary survey of noise levels in UK Secondary Schools, in Proceedings of the Institute of Acoustics and Belgium Acoustical Society Noise in the Built Environment, Gent, Belgium., April 29 - 30 2010.
- [14] H. Kuttruff, *Room acoustics*, Spon press, 2000.

- [15] A. Antoniou, and W.-S. Lu, *Practical Optimization: Algorithms and Engineering Applications* Springer, 2007.
- [16] A. Astolfi, V. Corrado, and A. Griginis: Comparison between measured and calculated parameters for the acoustical characterization of small classrooms, *Appl Acoust*, **69**,(11),(2008), 966-976.
- [17] A. Astolfi, V. Corrado, and M. Filippi, Classroom Acoustic Assessment: a procedure for field analysis, in Proceedings of the 5th European Conference on Noise Control, Naples, 2003
- [18] M. R. Hodgson: Experimental investigation of the acoustical characteristics of university classrooms, *J Acoust Soc Am*, **106**,(4),(1999), 1810–1819.
- [19] L. L. Beranek, *Acoustics*, McGraw-Hill, 1993.
- [20] N. Shiers, B. Shield, and R. Glanville, A survey of noise levels in a post-surgical children’s ward, in Proceedings of the Institute of Acoustics and Belgium Acoustical Society, Noise in the Built Environment, Gent, Belgium, April 29 - 30 2010.

Figure 1. Overview of the maximum likelihood blind acoustic parameter estimation algorithm.

Figure 2. Octave band  $RT_{20}$  for five university classrooms average parameter from six impulse responses measured using swept sine waves at three source locations and two receiver locations.

Figure 3. Source and receiver locations for a typical classroom measurement.

Figure 4. Error in blind  $RT_{20}$  estimation in noise free cases using 30 minutes of running speech for 5 different rooms, noise free condition calculated from one source-receiver position.

Figure 5. Largest error magnitude in blind  $RT_{20}$  estimate across all rooms, with one additional simultaneous talker as the noise source.

Figure 6. Largest error magnitude in blind  $RT_{20}$  estimate across all rooms, with six simultaneous talkers as the noise source.

Figure 7. Magnitude of bias error in blind  $RT_{20}$  estimates, verses estimated random error (95% confidence limits of blind  $RT_{20}$  estimates), where the noise source is one additional talker.

Figure 8. Magnitude of bias error in the blind  $RT_{20}$  estimates, verses estimated random error (95% confidence limits of blind  $RT_{20}$  estimates), where the noise source is six additional talkers.

Figure 9. Unoccupied (except for the researcher) octave band  $RT_{20}$  parameter for the classrooms, measured using 3 balloon bursts in each room.

Figure 10. Predicted  $RT_{20}$  verses occupancy level in the 1000 Hz octave band in Room A, with five occupied blind  $RT_{20}$  estimates and one unoccupied (except for the researcher)  $RT_{20}$  measurement. Solid line represents the model predicted  $RT_{20}$  vs occupancy level, while the two dotted lines represent the 95% confidence limits on that prediction.

Figure 11. Predicted  $RT_{20}$  verses occupancy level in the 1000 Hz octave band in Room B, with five occupied blind  $RT_{20}$  estimates and one unoccupied (except for the researcher)  $RT_{20}$  measurement. Solid line represents the model predicted  $RT_{20}$  vs occupancy level, while the two dotted lines represent the 95% confidence limits on that prediction.

Figure 12. Predicted  $RT_{20}$  verses occupancy level at the 1000 Hz octave band in Room C, with four occupied blind  $RT_{20}$  estimates and one unoccupied (except for the researcher)  $RT_{20}$  measurement. Solid line represents the model predicted  $RT_{20}$  vs occupancy level, while the two dotted lines represent the 95% confidence limits on that prediction.

Figure 13. Blindly estimated room absorption area,  $A_R$ , for Rooms A, B and C. Error bars show the 95% confidence limits on the estimates.

Figure 14. Comparison of absorption area per person,  $A_P$ , calculated using the blind estimation method with data from [16-19]. Error bars show the 95% confidence limits on the estimates.

Figure 15. Clinical skills laboratory at London South Bank University.

Figure 16.  $RT_{20}$  measurement of clinical skills lab, with and without the curtains drawn, measured using repeated balloon bursts.

Figure 17. Error in blind estimate of  $RT_{20}$ , for ward 2 for both the open and closed curtain case, in a simulated ward.

Figure 18. Blind  $RT_{20}$  estimates over a period of a week from Hospital 1 for five locations (day time data).

Figure 19. Blind  $RT_{20}$  estimates over a period of a week from Hospital 2 for five locations (day time data).

Figure 20. Blind  $RT_{20}$  estimates carried out in a four bed bay at hospital 1 before and after the replacement of ceiling tiles.

Figure 21. Comparison of day and night blind  $RT_{20}$  estimates for three hospital rooms.



Figure 1

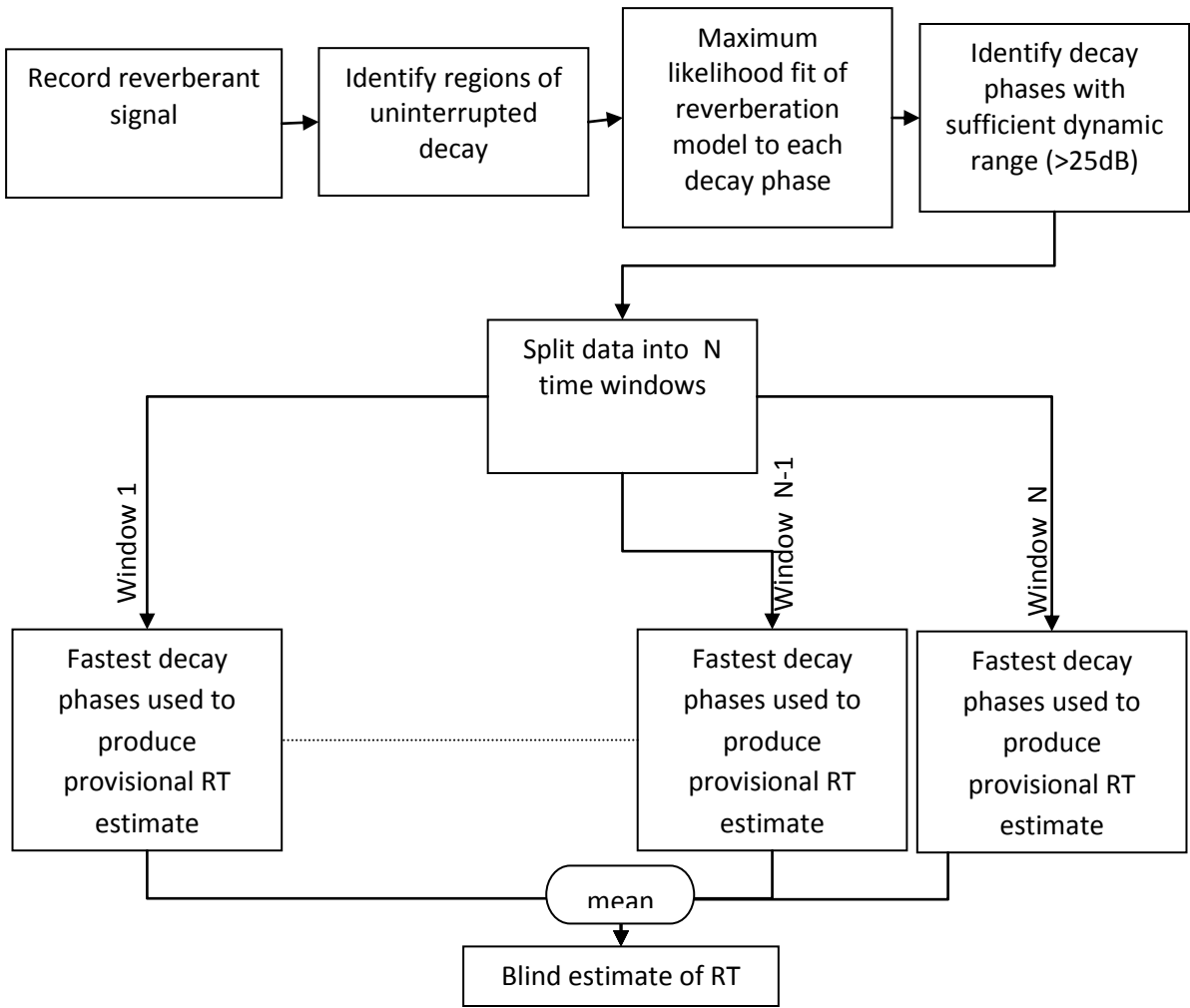


Figure 2

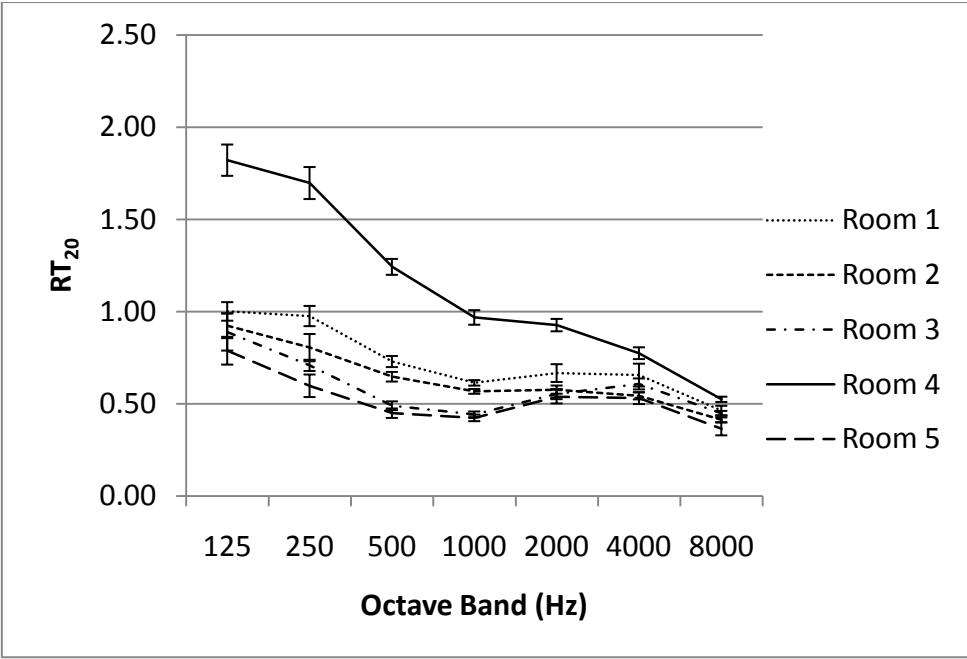


Figure 3

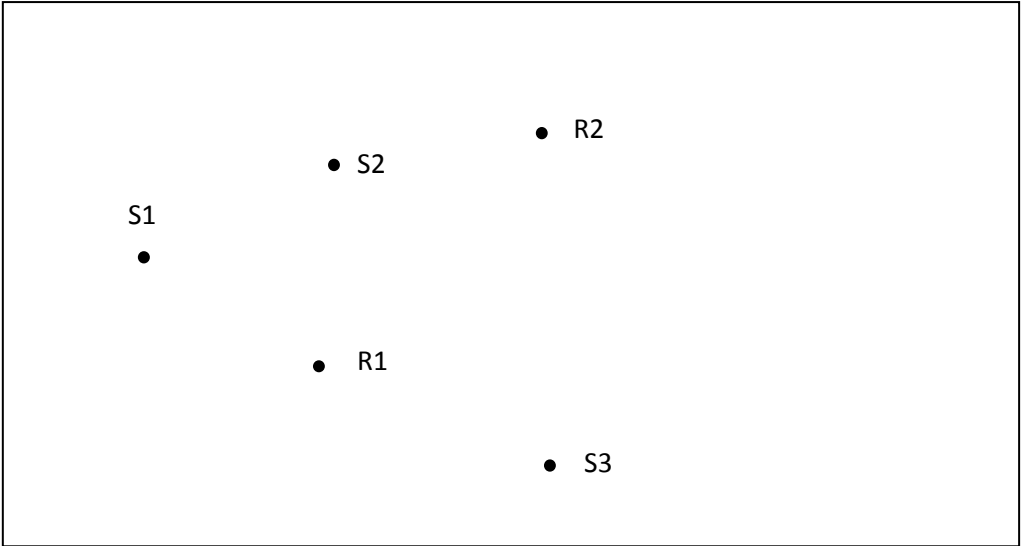


Figure 4

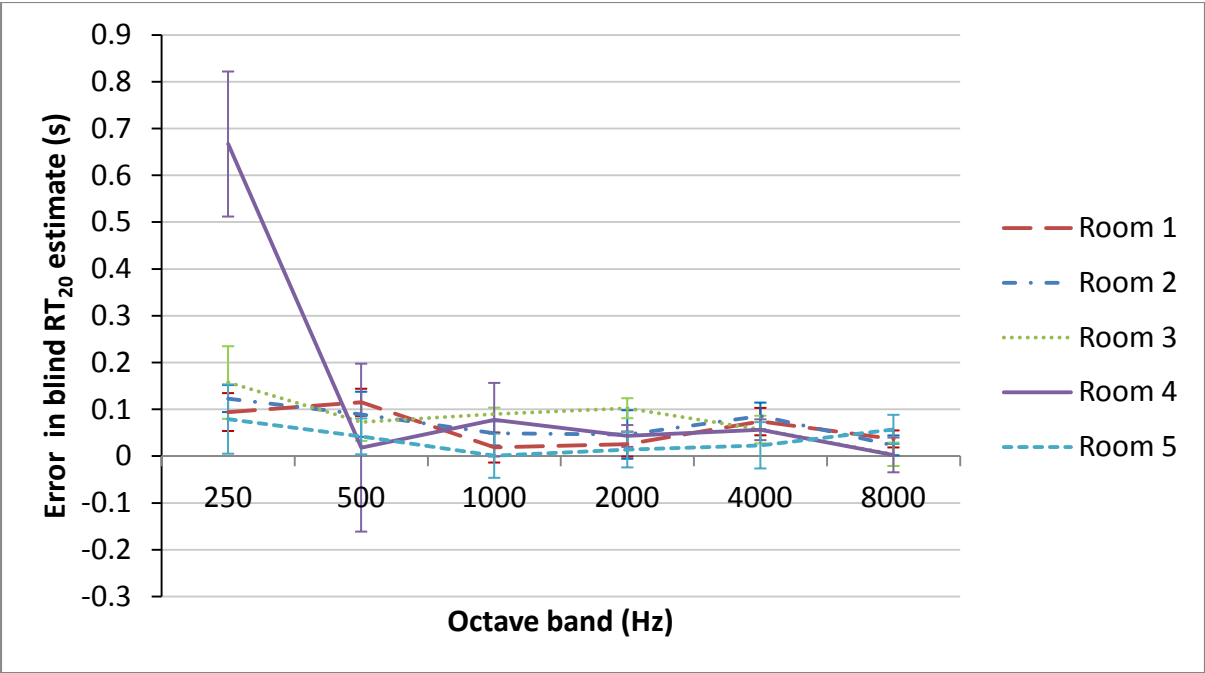


Figure 5

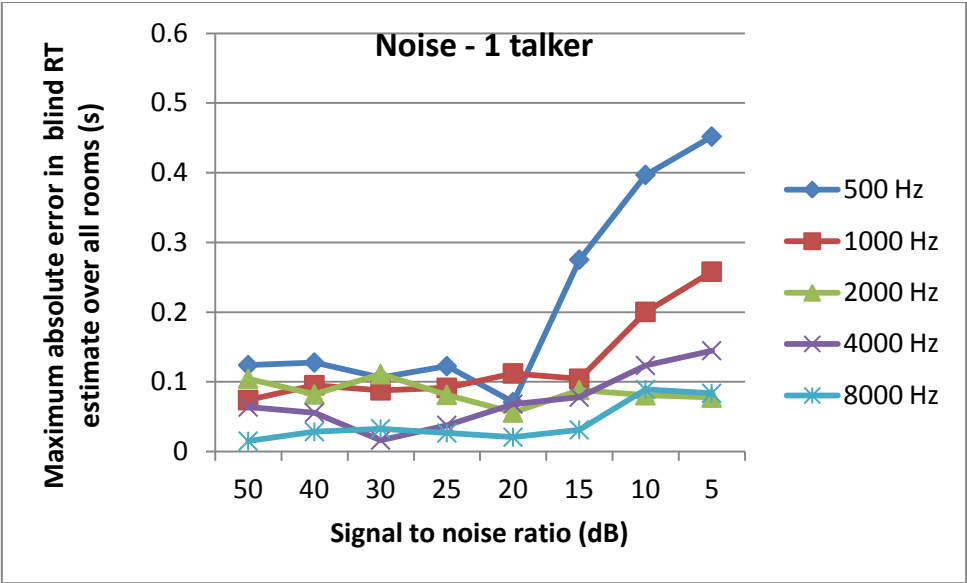


Figure 6

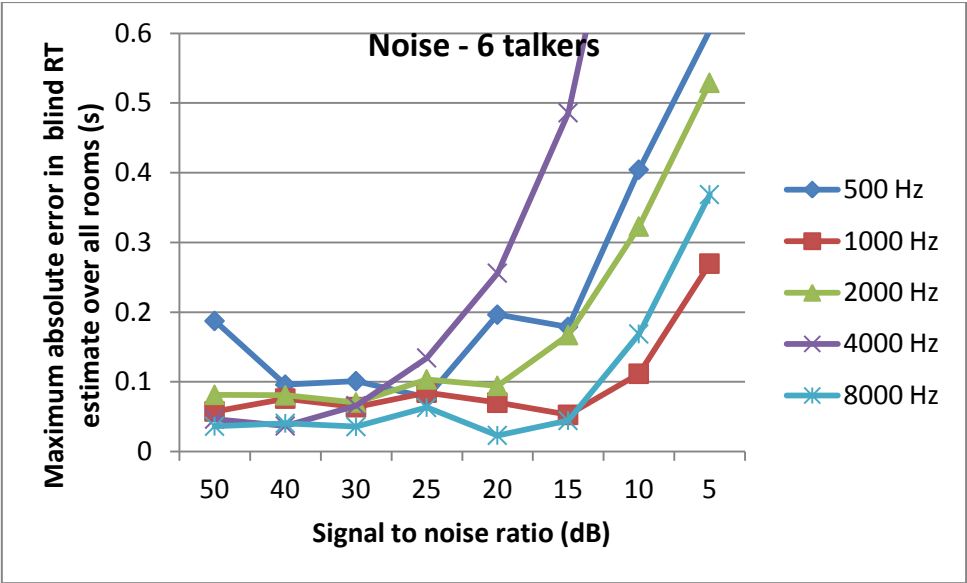


Figure 7

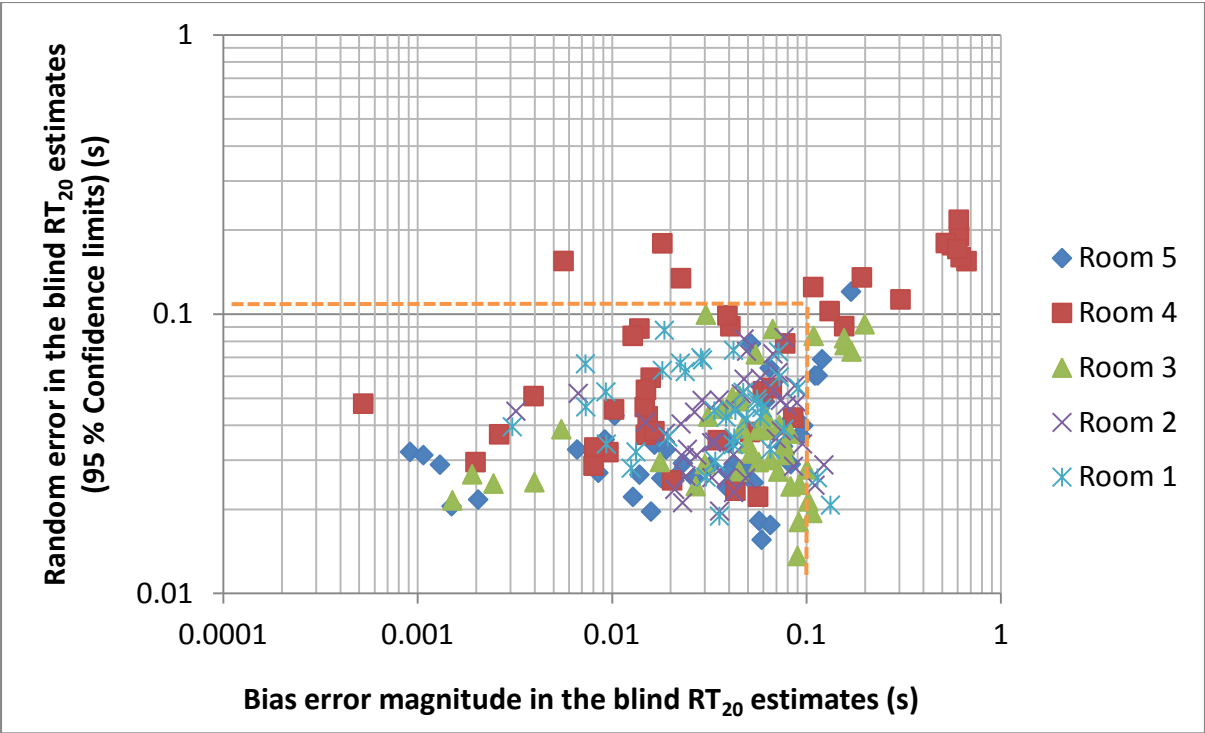


Figure 8

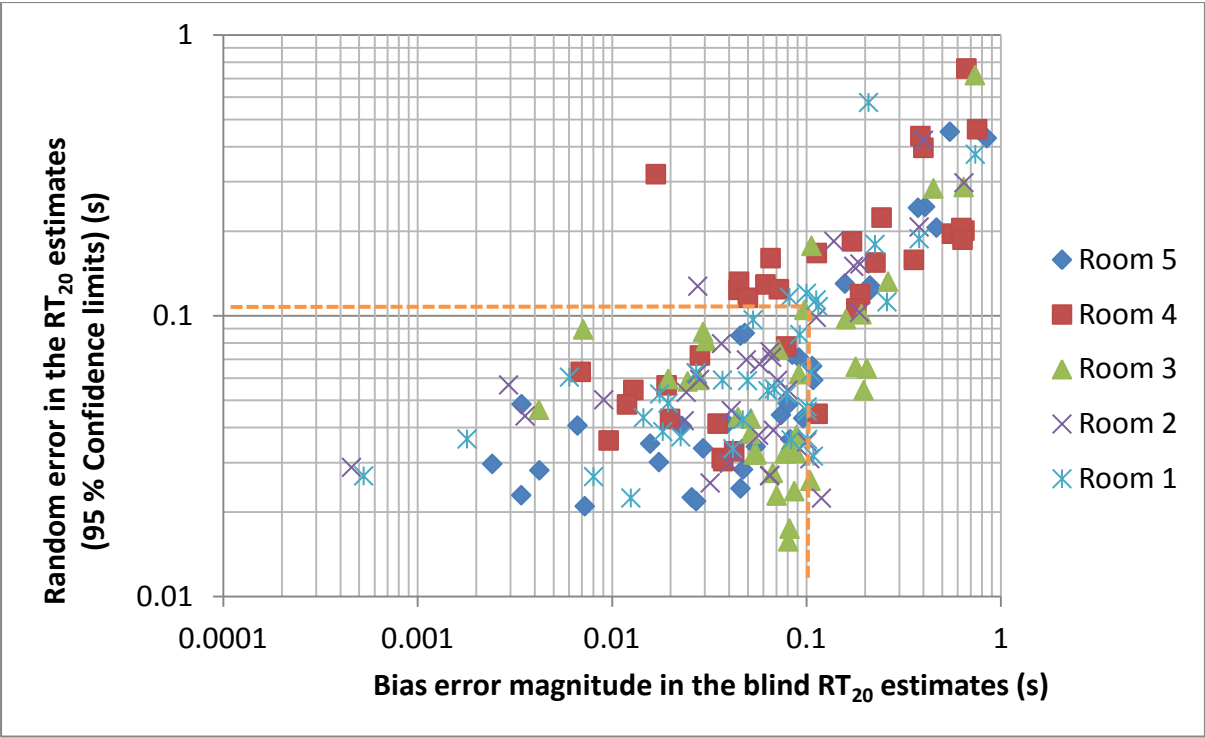




Figure 9

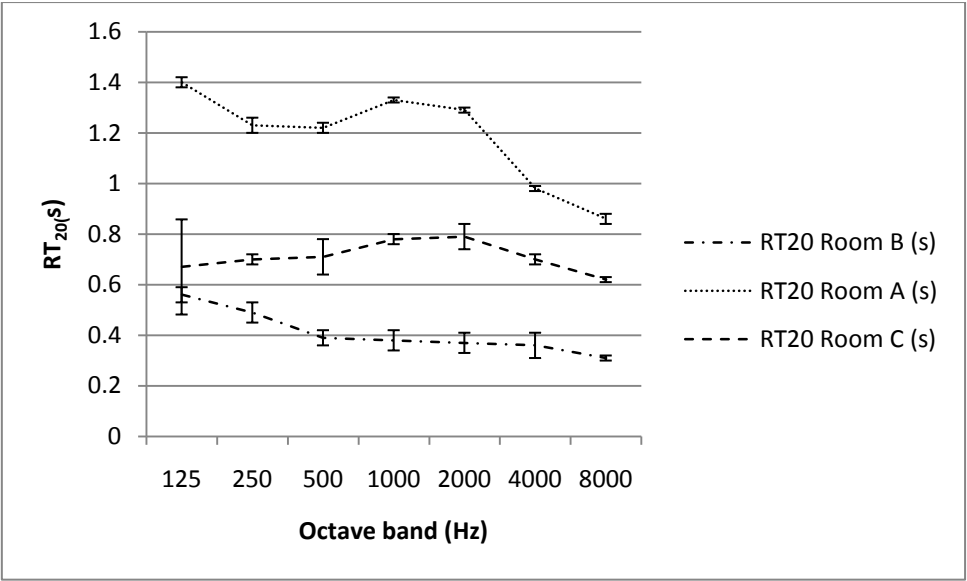


Figure 10

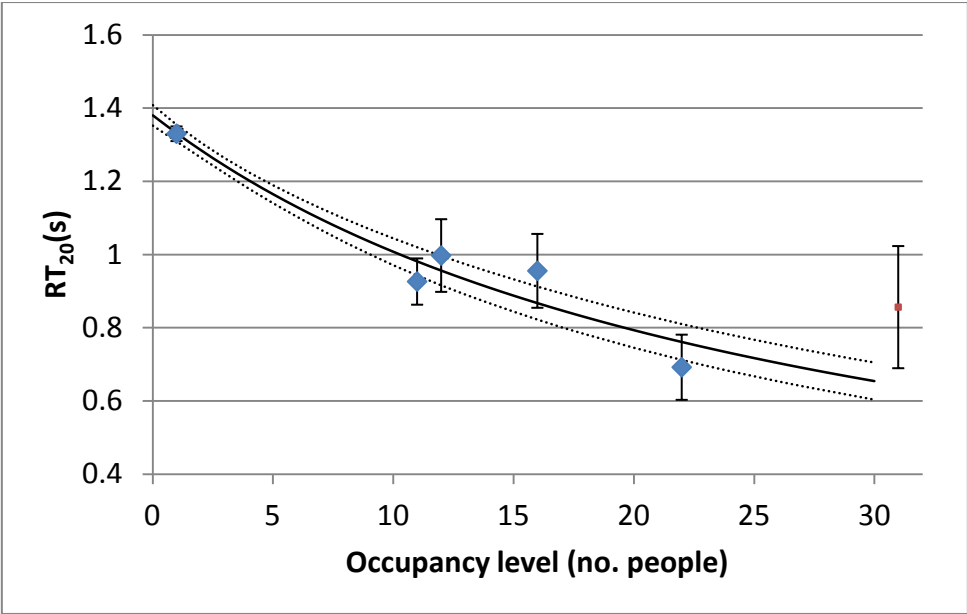


Figure 11

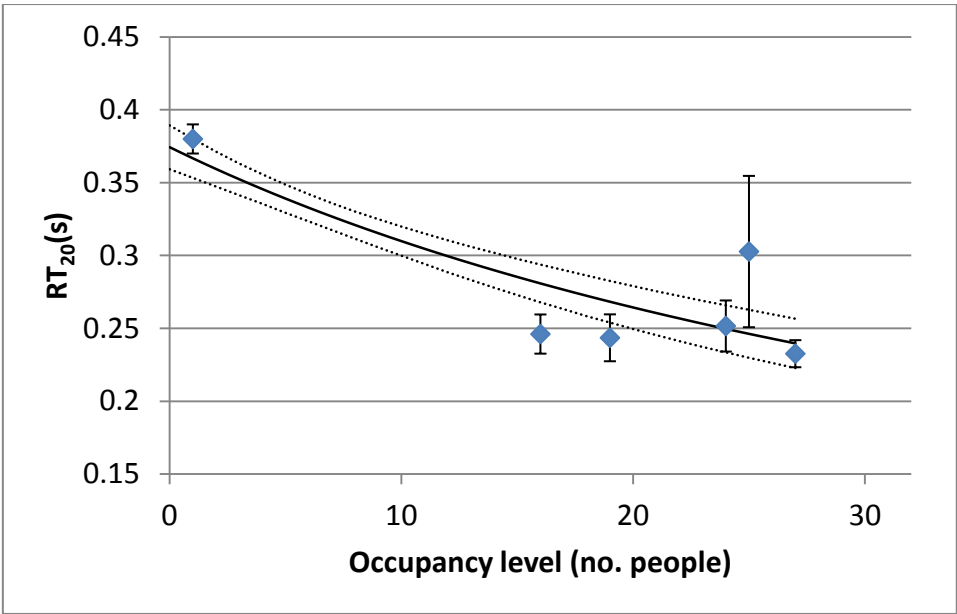


Figure 12

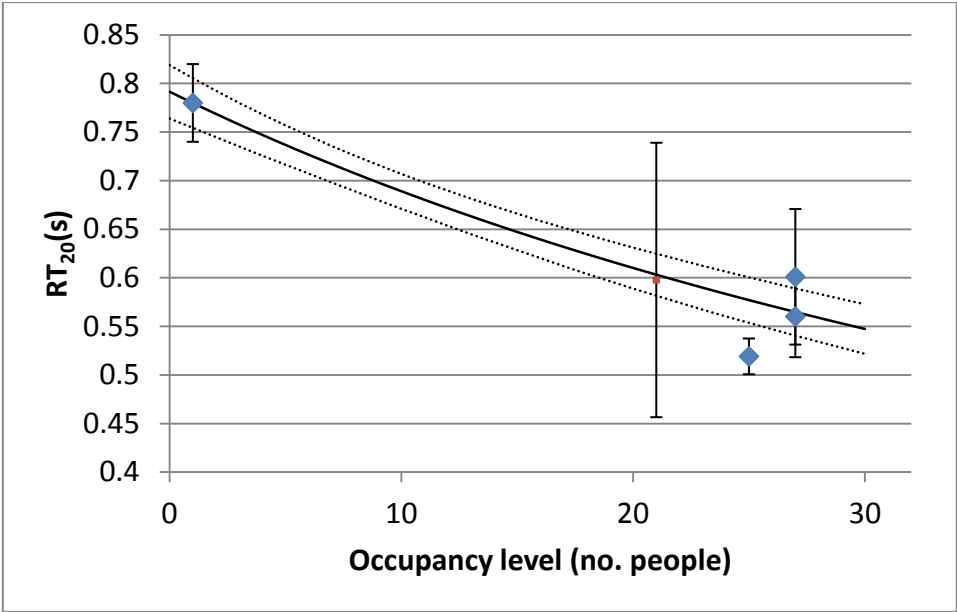


Figure 13

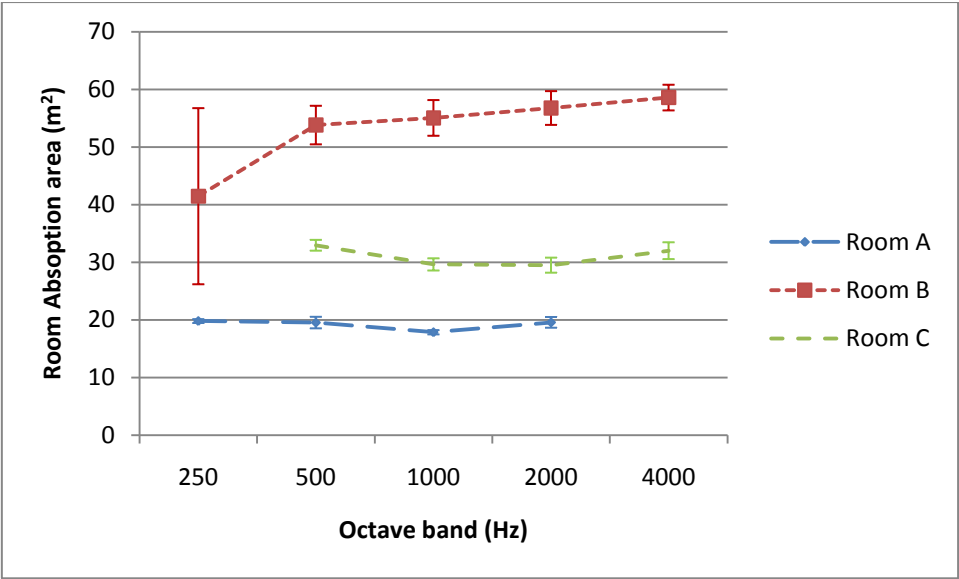


Figure 14

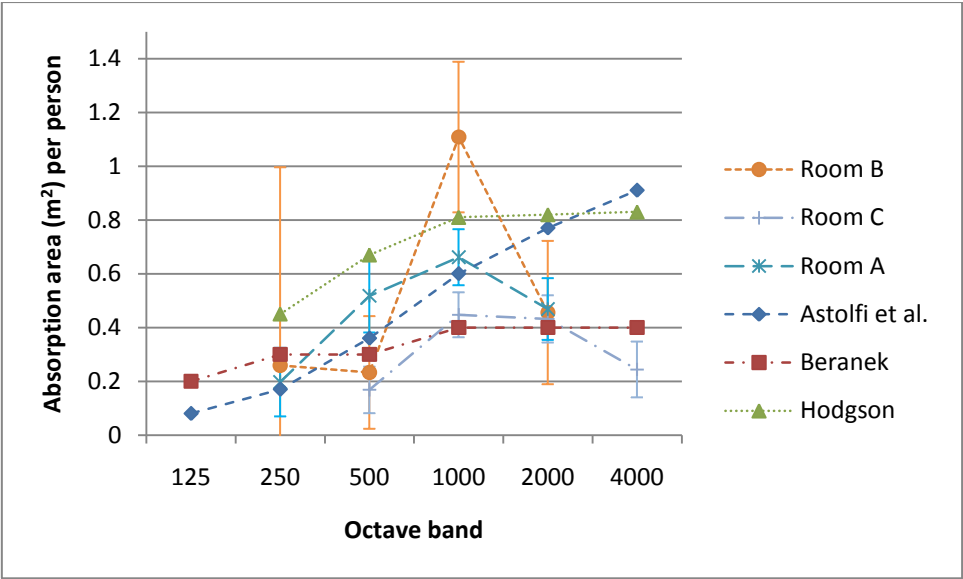


Figure 15



Figure 16

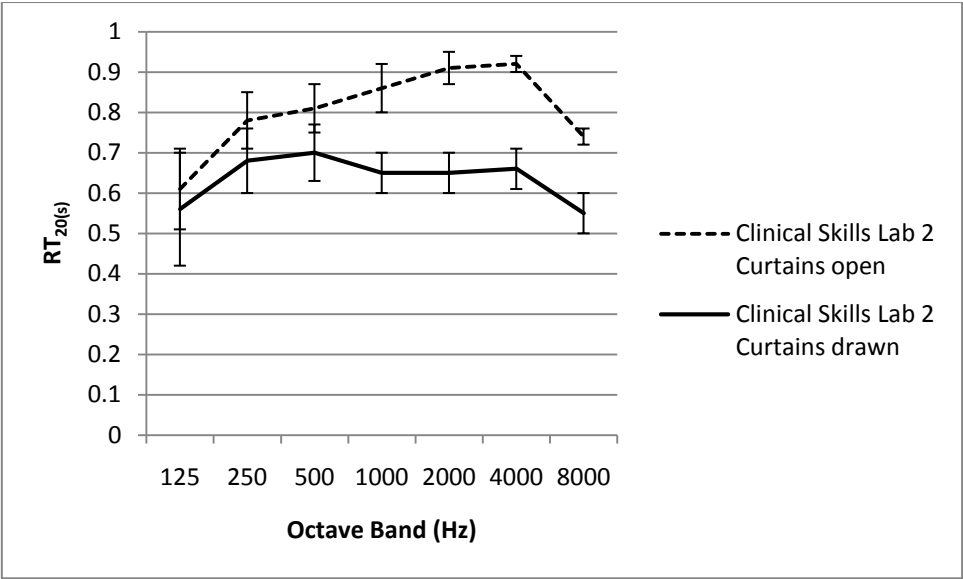




Figure 17

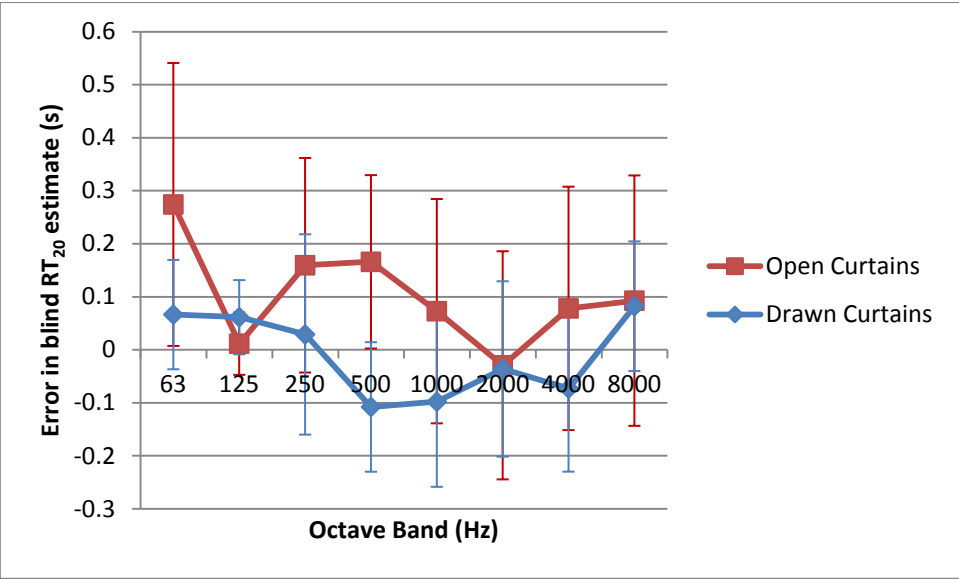


Figure 18

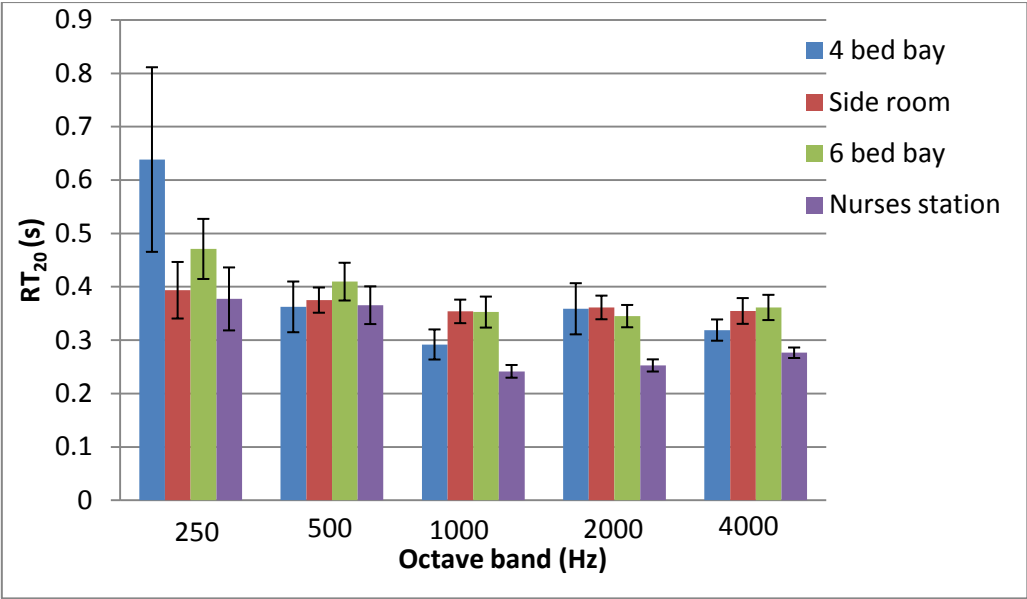


Figure 19

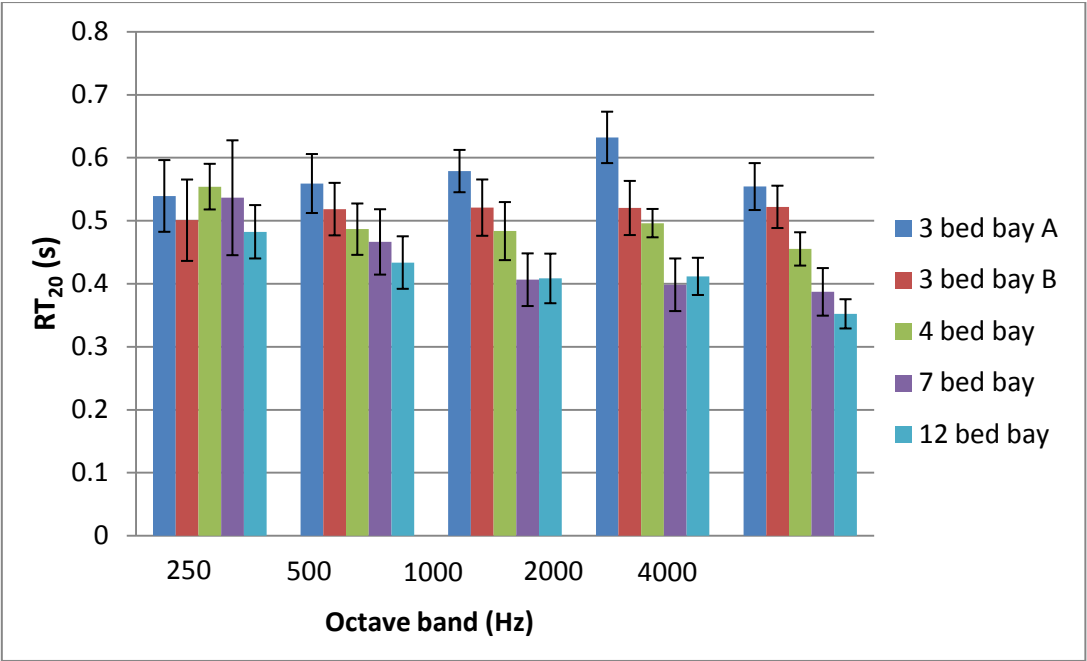


Figure 20

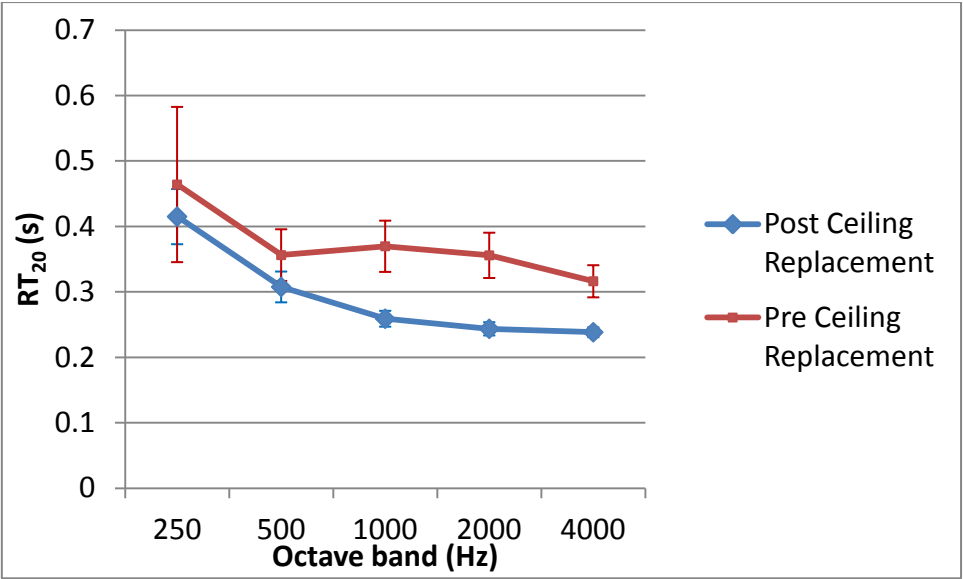


Figure 21

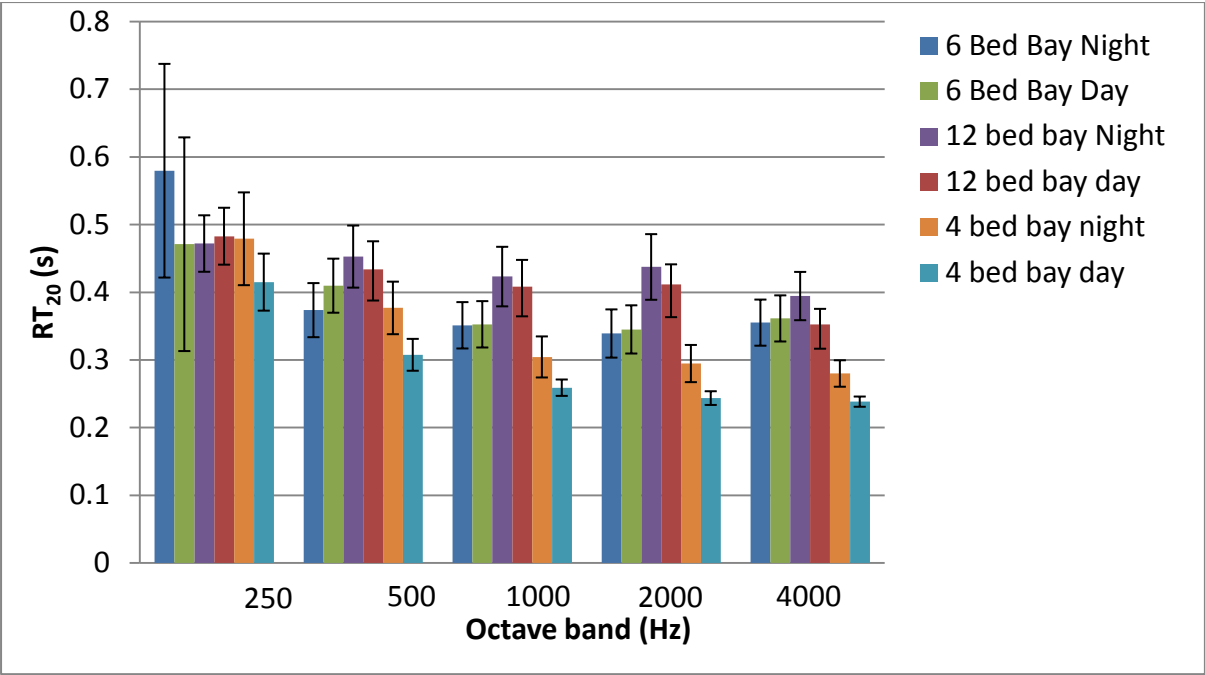


Table 1. 24 sound source mixtures, to be evaluated with the blind RT estimation algorithm

	<b>S1R1 convolved with...</b>	<b>+</b>	<b>S2R1 convolved with...</b>	<b>+</b>	<b>S3R1 convolved with...</b>
<b>Mixture 1</b>	Teacher sound source	+	Single talker competing source 1		
<b>Mixture 2</b>	Teacher sound source	+	Single talker competing source 1	+	Single talker competing source 2
<b>Mixture3</b>	Teacher sound source	+	3 talker competing source 1	+	3 talker competing source 1

Table 2. Occupancy levels for each lesson, and room volumes for each classroom

Lesson number	English Attendance	Maths Attendance	Science Attendance
L1	16	11	25
L2	25	12	27
L3	19	31	21
L4	24	16	27
L5	27	26	27
Room Volume	127m <sup>3</sup>	153m <sup>3</sup>	146m <sup>3</sup>

Table 3. Maximum 95% confidence limits on the predicted RT<sub>20</sub> over the full range of occupancies measured.

Octave band (Hz)	Room A (s)	Room B (s)	Room C (s)
250	± 0.14	± 0.04	N/A
500	± 0.07	± 0.03	± 0.04
1000	± 0.05	± 0.02	± 0.03
2000	± 0.06	± 0.03	± 0.04
4000	N/A	± 0.02	± 0.04

**Supplementary Material for on-line publication only**

**[Click here to download Supplementary Material for on-line publication only: Paul\\_kendrick\\_Blind estimation of reverberation time](#)**